

**PRIORITY  
DOCUMENT**  
SUBMITTED OR TRANSMITTED IN  
COMPLIANCE WITH RULE 17.1(a) OR (b)



EP04/51882

REC'D 18 OCT 2004
WIPO PCT

**Prioritätsbescheinigung über die Einreichung  
einer Patentanmeldung**

**Aktenzeichen:** 103 41 336.7  
**Anmeldetag:** 08. September 2003  
**Anmelder/Inhaber:** Siemens Aktiengesellschaft,  
80333 München/DE  
**Bezeichnung:** Verfahren zur optimierten Deaktivierung von  
Inter-Domain Routen  
**IPC:** H 04 L 29/14

**Die angehefteten Stücke sind eine richtige und genaue Wiedergabe der ur-  
sprünglichen Unterlagen dieser Patentanmeldung.**

München, den 30. September 2004  
**Deutsches Patent- und Markenamt**  
**Der Präsident**  
Im Auftrag

Dzierzon

## Beschreibung

## 5 Verfahren zur optimierten Deaktivierung von Inter-Domain Routen

Die Erfindung betrifft ein Verfahren zur Außerbetriebnahme von Inter-Domain Routen.

10 Die Erfindung liegt auf dem Gebiet der Internet-Technologien bzw. spezifischer auf dem Gebiet der Routing-Verfahren in paketorientierten Netzen.

15 Die derzeit wohl wichtigste Entwicklung auf dem Gebiet der Netze ist die Konvergenz von Sprach- und Datennetzen. Ein wichtiges Zukunftsszenario ist, dass Daten, Sprache und Videoinformationen über ein paketorientiertes Netz übertragen werden, wobei neuentwickelte Netztechnologien die Einhaltung von Anforderungsmerkmalen für verschiedene Verkehrsklassen gewährleisten. Die zukünftigen Netze für verschiedene Arten  
20 von Verkehr werden paketorientiert arbeiten. Aktuelle Entwicklungstätigkeiten betreffen die Übertragung von Sprachinformationen über herkömmlich für Datenverkehr verwendete Netze, vor allem IP (Internet Protokoll) basierte Netze.

25 Um Sprachkommunikation über Paketnetze und insbesondere IP-basierte Netze in einer Qualität zu ermöglichen, die der Sprachübertragung über leitungsvermittelte Netze entspricht, müssen Qualitätsparameter wie z.B. die Verzögerung von Datenpaketen oder der Jitter in engen Grenzen gehalten werden. Bei  
30 Sprachübertragung ist von großer Bedeutung für die Qualität des angebotenen Dienstes, dass die Verzögerungszeiten Werte von 150 Millisekunden nicht wesentlich übersteigen. Um eine entsprechend geringe Verzögerung zu erreichen, wird an verbesserten Routern und Routing-Algorithmen gearbeitet, die  
35 eine schnellere Bearbeitung der Datenpakete ermöglichen sollen.

Beim Routing über IP-Netze wird üblicherweise zwischen Intra-domain- und Interdomain-Routing unterschieden. Bei einer Datenübertragung über das Internet sind üblicherweise verschiedene Netze - man spricht häufig von Teilnetzen, von Domänen oder sogenannten autonomen Systemen (vom Englischen Autonomous System) - verschiedener Netzbetreiber involviert. Im Folgenden werden die Begriffe autonomes System und Inter-Domain Routing für das Routing zwischen autonomen Systemen verwendet. Routen entlang autonomen Systemen - auch Inter-Domain Routen genannte - werden im Folgenden auch einfach als Routen bezeichnet.

Die Netzbetreiber sind zuständig für das Routing innerhalb der autonomen Systeme, die in ihren Zuständigkeitsbereich fallen. Innerhalb dieser autonomen Systeme haben sie die Freiheit, die Vorgehensweise beim Routing nach eigenen Wünschen beliebig anzupassen, solange nur Dienstgütemerkmale eingehalten werden können. Anders stellt sich die Situation beim Inter-Domain Routing, d.h. beim Routing zwischen verschiedenen Domänen dar, bei dem unterschiedliche Betreiber von autonomen Systemen miteinander in Verbindung treten. Inter-Domain Routing wird dadurch verkompliziert, dass einerseits möglichst optimale Pfade über verschiedene autonome Systeme zu dem Ziel bestimmt werden sollen, andererseits aber die jeweiligen Betreiber lokal Strategien anwenden können, die eine globale Berechnung von optimalen Pfaden nach objektiven Kriterien erschweren.

Für das Routing zwischen verschiedenen Domänen werden sogenannten Exterior-Gateway-Protokolls EGP verwendet. Im Internet wird derzeit meist das in dem RFC (Request for Comments) 1771 genauer beschriebene Border-Gateway-Protokoll Version 4 (Border-Gateway-Protokoll wird häufig mit BGP abgekürzt) verwendet. Das Border-Gateway-Protokoll ist ein sogenanntes Path-Vector-Protokoll. Eine BGP-Instanz (der Ausdruck "BGP-Speaker" findet sich häufig in englischsprachigen Literatur) wird von seinen BGP-Nachbarn über mögliche Wege zu über den

jeweiligen BGP-Nachbar zu erreichenden Zielen informiert. Anhand von ebenfalls mitgeteilten Eigenschaften der Wege (im Englischen path attributes) bestimmt die BGP-Instanz zu den erreichbaren Zielen den aus ihrer lokalen Sicht jeweils optimalen Weg. Im Rahmen des BGP-Protokolls werden zwischen BGP-Instanzen vier Typen von Nachrichten ausgetauscht, darunter eine sogenannte UPDATE- oder Aktualisierungsnachricht, mit der Weginformationen durch das ganze Netz propagiert werden und die erlaubt, das Routing entsprechend der Erreichbarkeits-Änderungen zu optimieren. Die Aussendung von Aktualisierungsnachrichten führt üblicherweise zu einer Anpassung der Pfadinformationen bei allen betroffenen BGP-Instanzen des Netzes im Sinne eines entsprechend der lokal vorliegenden Informationen optimierten Routings. Daneben spielen sogenannte Keepalive- oder Zustandsbestätigungsnachrichten eine Rolle, mit denen eine BGP-Instanz seinen BGP-Nachbarn über seine Funktionsfähigkeit aufklärt. Bei Ausbleiben dieser Nachrichten gehen die BGP-Nachbarn davon aus, dass der Link zu der BGP-Instanz gestört ist.

Die Propagation von Erreichbarkeits-Informationen mit Hilfe des BGP-Protokolls hat den Nachteil, dass bei häufigen Änderungsanzeigen eine beträchtliche Last der durch das Netz propagierten Nachrichten zur Anzeige der Änderung auftreten und das Routing verhältnismäßig langsam konvergiert, d.h. dass die Umsetzung von Änderungen der Topologie des Netzes in entsprechende Inter-Domain Routen mit signifikanten Konvergenzzeiten verbunden ist.

Die Erfindung hat zur Aufgabe, ein Verfahren zur optimierten Fehlerbehebung beim Inter-Domain Routing anzugeben.

Die Aufgabe wird durch ein Verfahren nach Anspruch 1 gelöst.

Entsprechend der Erfindung wird der Ausfall eines zwei autonome Systeme verbindendes Segment (z.B. ein Link zwischen zwei jeweils einem der autonomen Systeme zugeordneten Rou-

tern) von einem Router (z.B. einem BGP speaker) eines ersten autonomen System an ein zweites autonomes System kommuniziert. Das zweite autonome System nimmt daraufhin die Routen außer Betrieb, welche das ausgefallene Segment enthalten. Eine Information über den Ausfall eines Segments kann zu allen betroffenen autonomen Systemen propagiert werden, welche dann die entsprechenden Routen deaktivieren. Bei Wiederinbetriebnahme des Segments kann eine entsprechende Nachricht an die autonomen Systeme gesendet werden, welche Routen deaktiviert haben und so eine Reaktivierung der Routen veranlasst werden.

Gemäß dem derzeitigen Stand der Technik, der durch das BGP Protokoll gegeben ist, werden Routen mittels UPDATE Nachrichten bekannt gegeben und im Fehlerfall durch sogenannte "withdrawals" explizit zurückgezogen (Routen im 'WITHDRAWN ROUTES'-Feld von UPDATE Nachrichten) oder durch die Bekanntgabe einer neuen Route zum selben Ziel implizit zurückgezogen. Eine Route besteht dabei aus zwei Teilen, einer Beschreibung der damit erreichbaren IP-Adressen (IP: Internet Protokoll) und einer Liste von AS-Nummern die in der Regel den AS-Pfad zum Zielnetz beschreibt. Dabei ist ein AS ein autonomes System. Im Fall eines fehlerhaften Segments in einer Route, beispielsweise einer fehlerhaften Verbindung zwischen zwei IP-Netzen, müssen alle darüber laufenden Routen explizit mittels withdrawals zurückgezogen werden oder durch die Bekanntgabe einer neuen Route implizit zurückgezogen werden. Meistens werden ausgehend von den an die fehlerhafte Verbindung angrenzenden Netzen explizite withdrawals von BGP-Instanz zu BGP-Instanz propagiert. Wenn eine BGP-Instanz eines weiter entfernt liegenden Netzes eine UPDATE-Nachricht mit einem withdrawal erhält und alternative Routen zum selben Ziel kennt, dann wählt sie eine diese Alternativen als neue Route und propagiert diese an Stelle des withdrawals. Da die BGP-Instanz jedoch die Ursache für das withdrawal nicht kennt, wählt sie häufig eine neue Route die sich zwar von der Zurückgezogenen unterscheidet, aber ebenfalls das fehlerhafte Segment enthält. Dadurch wird auf fehlerhafte Routen umge-

schaltet, bis durch weite UPDATE-Nachrichten auch diese Routen zurückgezogen werden. Zusammen mit dem MRAI (Minimum Route Advertisement Interval), einem Timer, der regelt, wie schnell Route-Updates weitergegeben werden, führt das Umschalten auf fehlerhafte Routen zu langen Konvergenzzeiten. Statt dem expliziten Zurückziehen aller betroffenen Routen im Fall einer fehlerhaften Verbindung zwischen zwei IP-Netzen wird gemäß der Erfindung das fehlerhafte Segment der betroffenen Routen unter den betroffenen Routern bzw. autonomen Systemen bekannt gegeben. Es ist bekannt, dass die Konvergenz bis zu 15 Minuten betragen kann, da im Fall der Zurückziehung einer Route nicht bekannt ist, wo das Problem liegt, und häufig zunächst auf bekannte alternative Routen umgeschaltet wird, die dasselbe fehlerhafte Segment enthalten.

Die neue Lösung führt zu wesentlich kürzeren Konvergenzzeiten. Wenn zusätzlich dafür Sorge getragen wurde, dass zu jedem Ziel-AS mindestens zwei verschiedene disjunkte Wege bekannt sind, kann damit die Verfügbarkeit wesentlich verbessert werden.

Weiter verringert sich durch die erfindungsgemäße Lösung die Menge an Information, die in einem Fehlerfall weiterzugeben ist erheblich: nur noch eine Nachricht statt viele Routen. Auch die Menge an auszutauschender Routinginformation stellt heute, wegen der Größe der Routingtabellen, ein ernstzunehmendes Problem dar. Dadurch, dass mit "Segment Withdrawals" wesentlich weniger Informationen auszutauschen ist, können diese effizienter und schneller verbreitet werden.

Die Erfindung hat zudem den Vorteil, dass sie mittels der durch herkömmliche Protokolle (namentlich dem BGP Protokoll) bereitgestellten Mittel realisiert werden kann. Dazu wird dieses Segment wie bisher wie eine Route beschrieben, enthält jedoch nur zwei AS-Nummern. Im Fehlerfall wird dieses Segment wie ein withdrawal weitergegeben, jedoch mit einer Markierung das es sich um ein Segment und nicht um eine komplette Route

handelt, z.B. mittels eines neuen "Path Attributes", welches entsprechend der Requests for Comments [RFC2042] und [RFC1771] benützt wird. Der Empfänger eines solchen "Segment Withdrawals" bzw. der Zurückziehung eines Segments behandelt 5 daraufhin alle ihm bekannten Routen, die das fehlerhafte Segment enthalten, als hätte er dazu ein route withdrawal erhalten.

10 Weiterhin lassen sich, wenn das fehlerhafte Segment wieder verfügbar ist (bei Fehlern hinreichend kurzer Dauer) alle betroffenen Routen in ähnlicher Weise durch ein "(Segment Withdrawal) Withdrawal" bzw. eine Rücknahme der Zurückziehung des Segments wieder in Betrieb nehmen. Diese Rücknahme kann ebenfalls mittels einer UPDATE Nachricht erfolgen. Die zu- 15 rückgezogenen Routen können bis zum Ablauf eines Timers mit dem Hinweis "für kurze Zeit nicht verfügbar" in den Routingtabellen bleiben.

20 Die Erfindung wird im Folgenden im Rahmen eines Ausführungsbeispiels anhand einer Figur näher dargestellt. Die Figur zeigt beispielhaft einen Verbund von IP-Netzen: AS1, AS2, ..., AS7 sind sieben administrativ unabhängige IP-Netze, die über die dargestellten Border-Router R11, R21, R22, R23, R31, R32, R41, R42, R43, R51, R52, R53, R61 und R71 IP-Verkehr untereinander austauschen. So gelangt beispielsweise ein an AS1 25 angeschlossener Kund an eine WWW-Seite eines an AS7 angeschlossenen Servers über die durch Sequenz von autonomen Systemen gegebene Route (1, 2, 4, 5, 7).

30 Wenn bei dem in der Figur gezeigten Verbund von IP-Netzen, welcher z.B. einen Ausschnitt aus dem Internet darstellt, die Verbindung von AS 5 zu AS 4 ausfällt, dann können mit dem Segment-Withdrawal:

(4, 5)

35 folgende Routen von AS 1 zurückgezogen werden:  
(1, 2, 4, 5)

(1, 2, 4, 5, 6)

(1, 2, 4, 5, 7)

und ein zwischenzeitliches Umschalten durch AS 2 auf eine der folgenden Alternativen vermieden werden:

5 (1, 2, 3, 4, 5)

(1, 2, 3, 4, 5, 6)

(1, 2, 3, 4, 5, 7)

d.h. alle Routen die die fehlerhafte AS-Folge (4, 5) enthalten sind mit einer einzigen Nachricht zurückgezogen. Kennt AS  
10 2 eine weitere nicht in der Figur dargestellte Route zu AS 5,  
so würde AS 2 unmittelbar darauf umschalten.



## Patentansprüche

1. Verfahren zur Außerbetriebnahme von Inter-Domain Routen, bei dem
  - 5 - von einem Router (R43) eines ersten autonomen Systems (AS4) der Ausfall eines zwei autonome Systeme (AS4, AS5) verbindenden Segments festgestellt wird,
  - von dem Router (R43) eine Nachricht an ein zweites autonomes System (AS3) gesendet wird, welche eine den Ausfall des
  - 10 Segmentes betreffende Information enthält, und
  - dass von einem Router (R32) des zweiten autonomen Systems (AS3) die das Segment enthaltende Inter-Domain Routen außer Betrieb genommen werden.
- 15 2. Verfahren nach Anspruch 1, dadurch gekennzeichnet, dass
  - die Nachricht von dem Router (R43) direkt oder über einen oder mehrere weitere Router (R42) an das zweite autonome Sys-
  - 20 tem gesendet wird.
3. Verfahren nach Anspruch 1 oder 2, dadurch gekennzeichnet, dass
  - der Router (R32) des zweiten autonomen Systems (AS3) sei-
  - 25 nerseits eine Nachricht, welche eine den Ausfall des Segmen-
  - tes betreffende Information enthält, zu mindestens einem wei-
  - teren, benachbarten autonomen System (AS2) sendet.
4. Verfahren nach Anspruch 3, dadurch gekennzeichnet, dass
  - 30 - die Nachricht von dem Router (R32) des zweiten autonomen Systems (AS3) direkt oder über einen oder mehrere weitere Router (R31) desselben autonomen Systems (AS3) an das weite-
  - re, benachbarte autonome System (AS2) gesendet wird.
- 35 5. Verfahren nach einem der vorhergehenden Ansprüche, dadurch gekennzeichnet, dass

- eine Nachricht, welche eine den Ausfall des Segmentes betreffende Information enthält, zu allen autonomen Systemen übermittelt wird, welche zumindest eine das Segment beinhaltende Route für das Routing von Datenpaketen aufweisen.

5

6. Verfahren nach einem der vorhergehenden Ansprüche, dadurch gekennzeichnet, dass

10 - bei Wiederinbetriebnahme des Segmentes eine Nachricht an ein autonomes System (AS3), welches das Segment enthaltende Inter-Domain Routen außer Betrieb genommen hat, gesendet wird, wobei die Nachricht eine Information darüber enthält, dass das Segment wieder in Betrieb genommen wurde.

15 7. Verfahren nach einem der vorhergehenden Ansprüche, dadurch gekennzeichnet, dass  
- bei Wiederinbetriebnahme des Segmentes eine Nachricht an alle autonomen Systeme, welche das Segment enthaltende Inter-Domain Routen außer Betrieb genommen haben, gesendet wird, wobei die Nachricht eine Information darüber enthält, dass  
20 das Segment wieder in Betrieb genommen wurde.

8. Verfahren nach Anspruch 6 oder 7, dadurch gekennzeichnet, dass  
25 - zumindest ein autonomes System (AS3), welches über die Wiederinbetriebnahme des Segmentes informiert wurde, die das Segment enthaltende Inter-Domain Routen wieder in Betrieb nimmt.

30 9. Verfahren nach einem der vorhergehenden Ansprüche, dadurch gekennzeichnet, dass  
- die Nachricht mit der Information über den Ausfall des Segments mittels einer UPDATE Nachricht des Border Gateway Protokolls übermittelt wird.

35 10. Verfahren nach einem der Ansprüche 6 bis 9, dadurch gekennzeichnet, dass

- die Nachricht mit der Information über die Wiederinbetriebnahme des Segments mittels einer UPDATE Nachricht des Border Gateway Protokolls übermittelt wird.

- 5 11. Verfahren nach Anspruch 9 oder 10,  
dadurch gekennzeichnet, dass
- das Segment in dem Feld der UPDATE Nachricht übermittelt wird, welches an sich für die Übermittlung von Routen vorgesehen ist, wobei mittels eines PATH ATTRIBUTES Parameters spezifiziert wird, dass es sich um ein Segment handelt.
- 10 12. Verfahren nach einem der vorhergehenden Ansprüche,  
dadurch gekennzeichnet, dass
- das erste und das zweite autonome System IP (internet Protocol) Netze sind.

## Zusammenfassung

## Verfahren zur optimierten Deaktivierung von Inter-Domain Routen

5

Das erfindungsgemäße Verfahren sieht vor, den Ausfall eines zwei autonome Systeme (AS4, AS5) verbindendes Segments als Segment-Ausfall betroffenen autonomen Systemen (AS3, AS4) zu kommunizieren. Die über den Segmentausfall informierten autonomen Systeme (AS3, AS4) deaktivieren daraufhin Inter-Domain Routen, die das ausgefallene Segment beinhalten. Nach Wiederinbetriebnahme des Segments kann nach Propagation einer entsprechenden Nachricht die das Segment umfassenden Routen wieder aktiviert werden. Das Verfahren kann herkömmliche Protokolle - namentlich das BGP Protokoll (Border Gateway Protokoll) - verwenden, indem das ausgefallene Segment als Route behandelt wird, welche mittels eines PATH ATTRIBUTES als Segment kenntlich gemacht wird. Im Gegensatz zum herkömmlichen Vorgehensweise im Rahmen des BGP Protokolls brauchen nicht

10

15

20

25

mehr verwendbare Routen nicht einzeln kommuniziert zu werden. Eine effizientere Propagation der Information über die Störung und eine bessere Konvergenz der daraufhin geänderten Routinginformationen der autonomen System wird damit erreicht.

Fig.

2003 13506  
Per

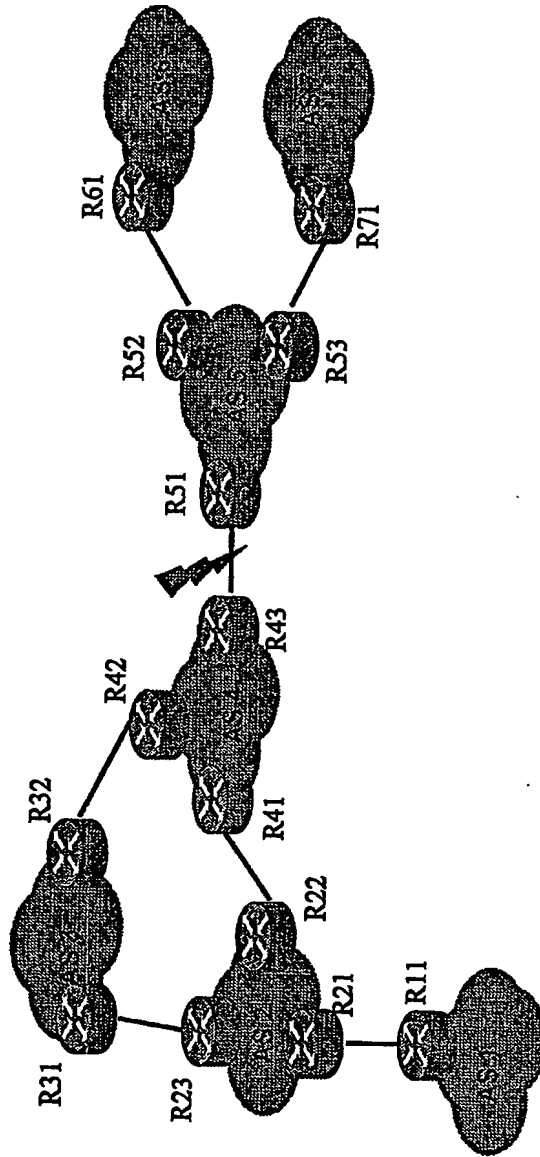


Fig.

BEST AVAILABLE COPY